



ELSEVIER

Centromeric chromatin: what makes it unique?

Steven Henikoff and Yamini Dalal

Centromeres represent the final frontier of eukaryotic genomes. Although they are defining features of chromosomes — the points at which spindle microtubules attach — the fundamental features that distinguish them from other parts of the chromosome remain mysterious. The function of centromeres is conserved throughout eukaryotic biology, but their DNA sequences are not. Rather, accumulating evidence favors chromatin-based centromeric identification. To understand how centromeric identity is maintained, researchers have studied DNA–protein interactions at native centromeres and ectopic ‘neocentromeres’. Other studies have taken a comparative approach focusing on centromere-specific proteins, of which mammalian CENP-A and CENP-C are the prototypes. Elucidating the assembly and structure of chromatin at centromeres remain key challenges.

Addresses

Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, PO Box 19024, Seattle, WA 98109-1024, USA

Corresponding author: Henikoff, Steven (steveh@fhcrc.org)

Current Opinion in Genetics & Development 2005, **15**:177–184

This review comes from a themed issue on
Chromosomes and expression mechanisms
Edited by Barbara Meyer and Jonathan Widom

Available online 11th February 2005

0959-437X/\$ – see front matter

© 2005 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.gde.2005.01.004

Introduction

Centromeres were described even before the rediscovery of Mendel’s laws of inheritance. By the 1880s, cytologists realized that chromosomes are characterized by a constriction that corresponds to the site of spindle attachment during mitosis [1]. As the only part of the chromosome that is directly involved in the process of mitosis, centromeres are fundamental to eukaryotic biology. Centromeric DNA attaches to spindle microtubules through a proteinaceous structure referred to as the kinetochore [2]. The outer kinetochore assembles when chromosomes condense and microtubules attach to it for mitosis, and it disassembles after the chromosomes have segregated. The inner kinetochore remains with the DNA throughout the cell cycle and consists of centromeric chromatin and associated ‘foundation’ proteins [3].

What has made centromeres so mysterious is the lack of any conserved sequence across species [4] despite the fact that all centromeres have an identical function — to organize the kinetochore at mitosis. This is not to say that centromeric DNA lacks any distinguishing features, because most centromeres comprise long stretches of short tandem repetitive ‘satellite’ DNA sequences that are found only there and in surrounding pericentric heterochromatin. The highly repetitive nature of centromeric satellites has made centromeres nearly intractable to sequence analysis except in special cases.

A key insight into the basis for centromere identity came from the discovery that a mammalian centromere-specific protein, centromere protein A (CENP-A), is homologous to histone H3 and is packaged into chromatin [5]. H3 is one of the four histones that form an octamer that packages the rest of the genome and that is assembled into nucleosomes during replication. Finding that H3 is replaced in centromeric chromatin by an H3-like variant suggests that centromeric chromatin is unique because of its histone complement and not because of its DNA sequence. Support for this hypothesis comes from the fact that CENP-A and its centromeric H3 (CenH3) counterparts in other organisms are found at all centromeres and are essential for centromere formation, whereas this is not the case for centromeric repeats [3].

There have been numerous reviews in recent years discussing diverse aspects of centromere structure, function and evolution [2,4,6–11]. In this review, we emphasize the progress that has been made in the understanding of centromeric chromatin since the subject was last reviewed in this series [9].

Centromere sequence organization

Centromeres in budding yeast are short and simple, and consist of common sequence elements that span just 125 bp [12]. One of these elements, CDEIII, is the binding site for the CDF3 multiprotein complex [13]. This complex is responsible for targeting the apparently single Cse4p-containing nucleosome — Cse4p is the yeast CenH3 [14]. Budding yeast centromeres, therefore, are well defined by DNA sequence alone.

Fission yeast centromeres are defined in a very different way than those of budding yeast. A central core of several kilobases that is rather dissimilar between chromosomes is surrounded by inverted ‘inner’ repeats, which are, in turn, surrounded by ‘outer’ repeats [9]. *Schizosaccharomyces pombe* CENP-A is found throughout the central core into the inner repeats surrounded by H3-containing

regions. There appears to be a remarkable separation of function for the central CENP-A-containing domain, which is where the kinetochore is thought to form, and the surrounding H3-containing domains, which are heterochromatic and are required for cohesion that is disrupted when centromeres separate at anaphase [15]. This model for centromeric chromatin organization appears to be general. In multicellular eukaryotes, the units are much larger and are apparently reiterated [16].

The small size and simplicity of budding yeast centromeres and the larger size and greater complexity of fission yeast centromeres might suggest that they form an evolutionary progression leading eventually to the megabase-sized centromeres of multicellular eukaryotes. This does not appear to be the case. A phylogenetic tree of life reveals that mammals and plants are each closer to the last common ancestor of eukaryotes, on the basis of sequence divergence, than are either budding or fission yeast (Figure 1a). In fact, these two groups of yeasts have diverged nearly as much from one another as have humans from *Arabidopsis*. The apparent simplicity of budding yeast centromeres, therefore, could easily have resulted from loss of a more complex centromeric organization, and it remains an open question as to what the ancestral centromere looked like [17]. This does not mean that budding yeast kinetochores are simple; rather, they are composed of more than 60 different proteins belonging to several distinct complexes that form a proteinaceous structure bigger than the ribosome [18,19]. The fact that only a minor fraction of these proteins appear to have mammalian counterparts might reflect the recruitment of new proteins down the budding yeast lineage.

Whereas fungal centromeres are diverse in structure, it is remarkable that plants and animals have nearly identical overall structures despite the fact that the divergence between plants and animals preceded that between animals and fungi [20]. Specifically, human centromeric satellites comprise several megabases of an AT-rich 171 bp α -satellite repeat [21], and *Arabidopsis* centromeric satellites comprise 3 Mb of a different AT-rich 178 bp satellite [22]. Despite these structural similarities, differences in satellite sequences are the rule, even between closely related species, indicating that centromeric satellites are the most rapidly evolving regions of complex genomes [4].

Important progress has been made in sequencing and assembling highly repetitive sequences, and we are witnessing the inexorable drive to fill centromeric gaps [23^{*},24^{*}]. Because of this progress, human centromeres are becoming useful model systems for understanding centromere organization. The functional CENP-A-containing region of a human centromere lies in almost perfectly homogeneous higher-order repeats of the basic 171 bp α -satellite repeats, surrounded by increasingly diverged α satellite, which becomes rich in long inter-

persed element 1 (LINE-1) elements near the flanks [25]. Although the 171 bp human repeat and the 178 bp *Arabidopsis* repeat are unrelated in sequence, both are peppered with an abundant retrotransposon near the edges [26]. The overall organization of plant and animal centromeres, therefore, appears to be very similar.

DNA-protein interactions

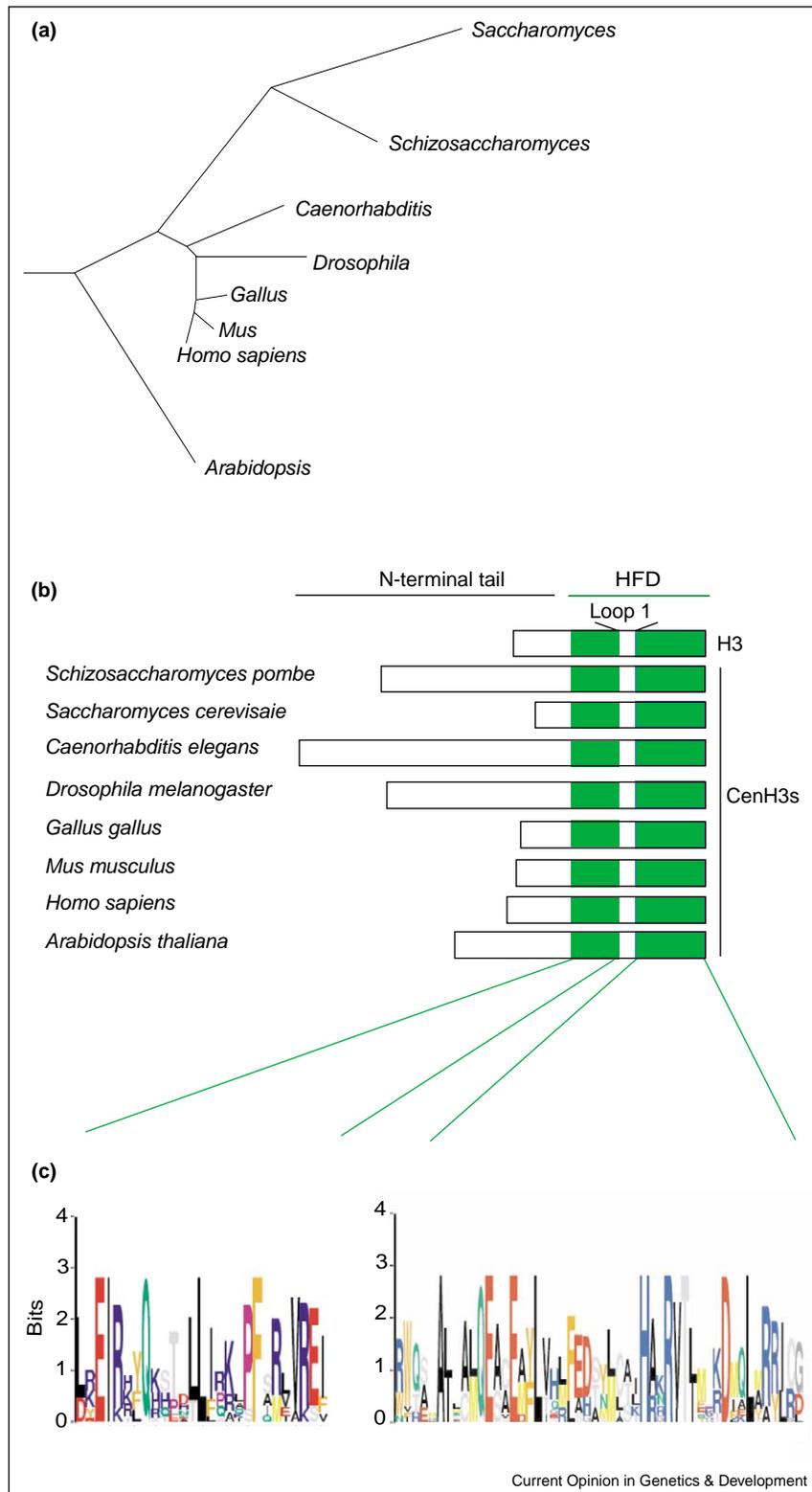
Although dozens of proteins localize to the kinetochore [18,19], the large majority of these are seen only at mitosis, and only two of these, CenH3 and centromere protein C (CENP-C), are known to bind DNA and to be widely distributed in evolution [8]. In every case that has been examined, CENP-C localization depends on the presence of CenH3 but not vice versa [27–31]. CenH3-containing nucleosomes appear, therefore, to provide the chromatin framework for centromeres, and much recent attention has been focused on the properties and evolution of CenH3s.

CENP-A-containing nucleosomes were first described many years ago [32] and have been successfully assembled *in vitro* [33]. Recently, soluble CENP-A–H4 tetramer particles were found to have properties *in vitro* that are suggestive of a more compact and rigid structure than that of H3–H4 tetramers [34^{*}]. The region responsible for this structural difference includes Loop 1 of the histone fold domain (Figures 1b,c), which had previously been shown to be both necessary and sufficient for centromere localization of the *Drosophila* CenH3, CID (centromere identifier) [35]. Further evidence that Loop 1, together with an adjacent region of the core, is involved in DNA-binding specificity comes from the discovery that these regions are adaptively evolving in *Drosophila* and *Arabidopsis* [36,37].

Many CenH3s, including mammalian CENP-A, are not adaptively evolving; however, in all plant and animal lineages examined, CENP-C is adaptively evolving over extensive stretches of the protein [38^{*}]. Adaptively evolving regions of CENP-C roughly correspond to sites of DNA binding and centromere targeting. By contrast, yeast Cse4p and CENP-C are under strict purifying selection, which is consistent with the idea that complex centromeres composed of rapidly evolving DNA interact with adaptively evolving components, whereas those with simple centromeres have evolutionarily fixed protein components.

The central role that CENP-C appears to play in centromere specificity in humans might help account for the surprising observation that budding yeast Cse4p not only localizes to human centromeres but can also functionally replace CENP-A [39^{**}]. Replacement was accomplished by extinguishing CENP-A mRNA while Cse4p was being produced, thus allowing Cse4p to fill in gaps created by CENP-A loss, while retaining all other centromeric com-

Figure 1



CenH3s align in the histone fold domain (HFD) but have dissimilar tails. **(a)** Evolutionary distances for selected key species based on a phylogenetic tree of life [20]. **(b)** Alignment of CenH3s from selected species. The amino termini are not conserved in length or sequence in different lineages. **(c)** Sequence LOGOS format shows variable conservation of the HFD.

ponents. It appears that such promiscuous replacement is lineage-specific because *Drosophila bipectinata* CID fails even to localize to *Drosophila melanogaster* centromeres, whereas chicken CENP-A is capable of localizing to mouse centromeric satellite DNA [35,40]. This difference might be attributable to absence of CENP-C in *Drosophila* species and the adaptive evolution of CID [38^{*}]. Taken together, these observations suggest that whereas localization of centromeric components can differ between lineages the resulting chromatin structure is likely to be universal.

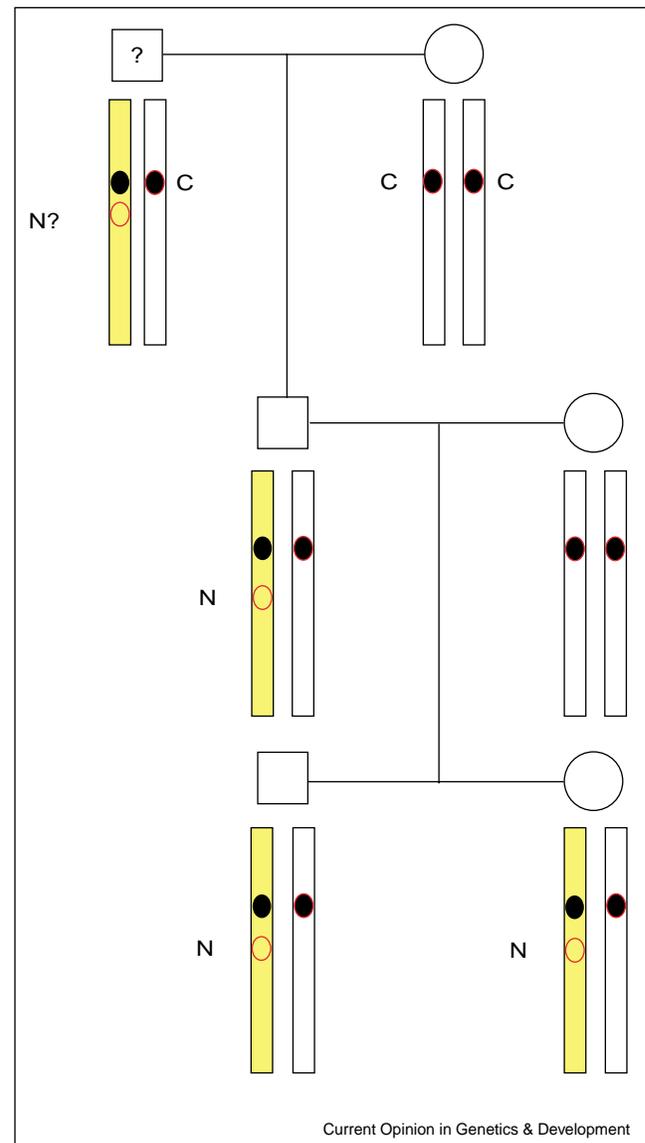
Centromere evolution

Despite evidence for specificity between centromere-binding proteins and centromeric DNA, no sequence determinants have been identified for any complex centromere. This conclusion comes primarily from the existence of ~70 different human neocentromeres that entirely lack α -satellite DNA, which is usually discovered from karyotype analysis that follows a diagnosis of possible aneuploidy [7]. Rarely, neocentromeres are discovered in otherwise karyotypically normal individuals, and three such cases have been documented [41,42^{*},43^{*}]. In all three cases, the neocentromere was inherited from parent to offspring and did not appear to be responsible for any abnormalities. One case involves a shift in centromere position within chromosome 4 (Figure 2), in which instance at least 1 Mb of α satellite was retained at the native centromere [42^{*}]. It is unknown whether the native centromere has suffered an inactivating genetic lesion or, alternatively, has been epigenetically silenced.

The existence of human neocentromeres provides the most compelling evidence that α -satellite DNA is not necessary for centromere function and that ordinary regions of chromosomes are competent to acquire centromeric function [7]. In general, these regions are LINE-rich and AT-rich and short interspersed element (SINE)-poor and gene-poor. The documentation of gene activity throughout one of these neocentromeres rules out a requirement for constitutive heterochromatin at centromeres [44^{**}]. One possibility is that genes and centromeres can cohabit the same region of DNA because transcription does not occur during mitosis, which is the only time during which centromeres function [45^{**}].

Cytological evidence suggested that neocentromeres occur at common positions [43^{*}]. Surprisingly, even neocentromeres that arose in the same cytological band occupy non-overlapping positions on a molecular scale [46^{**}]. One might also expect that neocentromeres would form preferentially at sites at which centromeres have formed in primate ancestors. Indeed, a site of formation of an Old World monkey centromere was deduced to coincide with the cytological position of human neocentromeres on chromosome 3 [43^{*}]. However, neither the human neocentromeres nor the monkey centromere

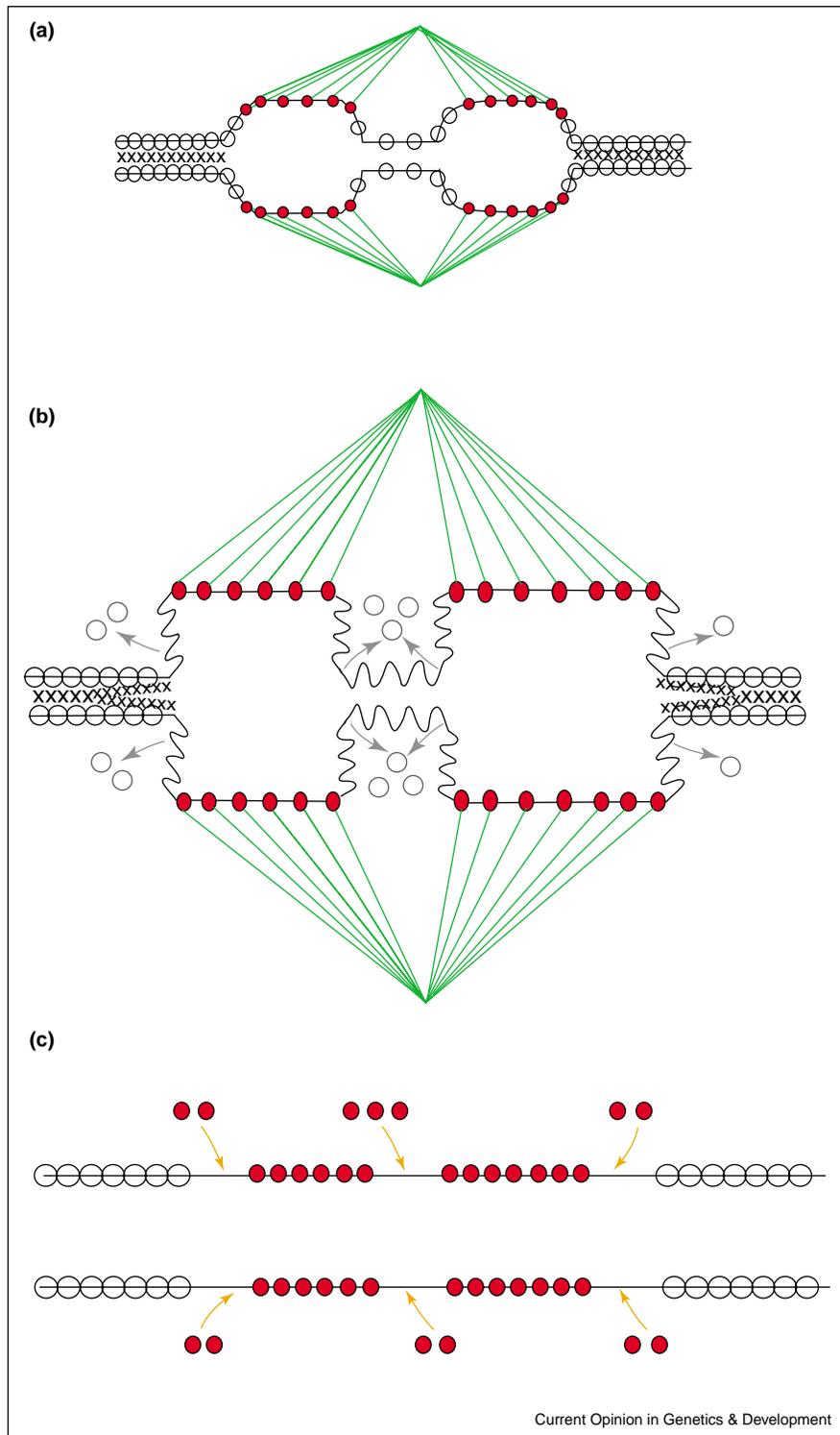
Figure 2



A case of human chromosome 4 repositioning. Pedigree showing that a chromosome 4 neocentromere (on the yellow chromosome) is inherited through at least two generations. The karyotype of the grandfather (?) was not determined. Abbreviations: C, chromosome 4 centromere; N, Neocentromere. Red open circles indicate the chromosomal location of the active centromere in each generation. Black filled circles indicate alpha satellite sequences at the original centromere. Based on [42^{*}].

occupy overlapping molecular positions. These observations suggest that a very large portion of the human genome is competent to form neocentromeres, but we are limited in our ability to discover them because centromere shifts are discovered fortuitously and other neocentromeres usually result in severe aneuploidy. Neocentromere formation might be much more frequent than their rate of discovery because only cases of viable aneuploidy would lead to their ready detection [7].

Figure 3



A model in which centromeric function reinforces centromeric chromatin integrity [9,51]. **(a)** Mitotic spindles (green lines) just before anaphase. Microtubules are bound exclusively to specialized centromeric chromatin, but not to H3 nucleosomes deposited within centromeric arrays during replication. Open circles are H3-containing nucleosomes, red circles are centromeric nucleosomes. Cohesion is indicated by X between H3 nucleosomes. **(b)** At anaphase, spindle microtubules pull on centromeric chromatin. Centromeric nucleosomes are resistant to unraveling; however, H3 nucleosomes unravel and are expelled from the array creating new gaps. **(c)** Gaps in the centromeric chromatin array are preferentially filled post-mitosis by centromeric nucleosomes.

If ordinary regions of the genome are competent to form neocentromeres and do so frequently over evolutionary time, why are all native centromeres composed of tandemly repeated satellite sequences? One possibility is that neocentromeres represent only the earliest stage in centromere evolution, and expansion of satellite sequence arrays occurs so rapidly that intermediates are never seen. If so, then a broad survey of species might reveal incipient centromeres that have only recently arisen. This evolutionary scenario has been proposed for rice [45**], whose 11 chromosomes contain a range of satellite arrays at their functional centromeres that can be as small as ~50 kb. The lack of extensive satellite arrays in the centromere of rice chromosome 8 has enabled its essentially complete sequence to be determined [45**,47]. The kinetochore region was localized to a 750 kb region by CenH3 binding and shown to contain active genes just as was found for a human neocentromere [44**,45**]. Furthermore, CenH3-binding appeared to overlap that of H3, which suggests cell-to-cell variability in the location of centromeric nucleosomes within the functional centromere. This finding supports a model of centromere plasticity inferred from an observed correlation between the relative extent of CenH3 and H3 nucleosomal arrays and the availability of the two histones [16].

Assembly of centromeric chromatin and kinetochore function

There seems to be little doubt that centromere identity and the presence of CenH3-containing nucleosomes are inseparable. It is still an open question, however, whether or not there are undiscovered centromere determinants, because overproduction of CENP-A in human cells recruited CENP-C but did not lead to ectopic centromere formation [30]. Furthermore, human artificial centromeres have been produced using amplified α -satellite repeats found at native centromeres but not from α -satellite repeats found in surrounding regions [25]. Results like this have discouraged attempts to create artificial centromere-containing plasmids for stably maintaining transgenes in plants and animals as was done long ago in yeast [48].

We suspect that a solution to the problem of how to create artificial centromeres might present itself if we only knew how centromeric nucleosomes are assembled in the same place during every cell cycle. Studies of tagged CenH3s show that they can deposit at centromeres in a replication-independent way [49,50], and that their overproduction causes them to be deposited throughout euchromatin [51,52]. We speculate that replication-independent deposition of centromeric nucleosomes is normally occurring both in centromeres and in euchromatin, but in euchromatin there is a normal replacement process that appears continually to deposit nucleosomes containing the histone H3.3 variant [51]. There might be, however, no such turnover at centromeres so that centromeric

nucleosomes are deposited at gaps in CenH3-containing arrays and are retained.

Replication-independent nucleosome assembly requires the creation of gaps between nucleosomes, but it is not clear how this occurs at centromeres. Gaps would occur behind the replication fork as old nucleosomes distribute between the two daughter strands if canonical nucleosomes are inhibited from assembling there [21]. Consistent with this possibility, CENP-A deposits at G_2 in human cells [50]. An intriguing additional possibility is that gaps also occur at anaphase from nucleosomes that have unraveled because of the tension exerted by spindle microtubules [9,51] (Figure 3). In this case, CenH3-containing nucleosomes must be retained because they are attached to the spindle microtubules and withstand the tension. The nearby H3-containing nucleosomes, however, are not attached so they are subject to unraveling and subsequent replacement by their CenH3-containing counterparts. Thus, a key feature of CenH3-containing nucleosomes would be that despite their attachment to the spindle they are so tightly bound to centromeric DNA that they can remain in place when they are pulled upon, whereas their H3-containing neighbors are unraveled. In this way, centromere function would reinforce centromere integrity, and plasticity would result from replacement of H3 nucleosomes adjacent to CenH3 nucleosomes at each mitosis.

Conclusions

Despite considerable recent progress in defining centromeric components, we still lack a clear understanding of how centromeres are distinguished from ordinary regions of the genome that are not centromeres. Maintenance of centromeric chromatin, which forms the foundation for spindle attachment, is extraordinary insofar as centromeres remain in the same cytological position over tens of millions of years. Yet the occasional appearance of neocentromeres in numerous regions of ordinary sequence composition and gene content indicates that centromeres are not constrained by any recognizable sequence dependence. With a better understanding of the process that assembles centromeric chromatin during the cell cycle, we might be able to solve this continuing puzzle.

Acknowledgements

We thank Takè Furuyama and Paul Talbert for comments on the manuscript.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Flemming W: *Zellsubstanz, Kern und Zelltheilung*. Leipzig: FCW Vogel; 1882.
2. Cleveland DW, Mao Y, Sullivan KF: **Centromeres and kinetochores: from epigenetics to mitotic checkpoint signalling**. *Cell* 2003, **112**:407-421.

3. Choo KH: **Domain organization at the centromere and neocentromere.** *Dev Cell* 2001, **1**:165-177.
 4. Malik HS, Henikoff S: **Conflict begets complexity: the evolution of centromeres.** *Curr Opin Genet Dev* 2002, **12**:711-718.
 5. Palmer DK, O'Day K, Trong HL, Charbonneau H, Margolis RL: **Purification of the centromere-specific protein CENP-A and demonstration that it is a distinctive histone.** *Proc Natl Acad Sci USA* 1991, **88**:3734-3738.
 6. Hall AE, Keith KC, Hall SE, Copenhaver GP, Preuss D: **The rapidly evolving field of plant centromeres.** *Curr Opin Plant Biol* 2004, **7**:108-114.
 7. Warburton PE: **Chromosomal dynamics of human neocentromere formation.** *Chromosome Res* 2004, **12**:617-626.
 8. Amor DJ, Kalitsis P, Sumer H, Choo KH: **Building the centromere: from foundation proteins to 3D organization.** *Trends Cell Biol* 2004, **14**:359-368.
 9. Mellone BG, Allshire RC: **Stretching it: putting the CEN(P-A) in centromere.** *Curr Opin Genet Dev* 2003, **13**:191-198.
 10. Fukagawa T: **Assembly of kinetochores in vertebrate cells.** *Exp Cell Res* 2004, **296**:21-27.
 11. Jiang J, Birchler JA, Parrott WA, Dawe RK: **A molecular view of plant centromeres.** *Trends Plant Sci* 2003, **8**:570-575.
 12. Fitzgerald-Hayes M, Clarke L, Carbon J: **Nucleotide sequence comparisons and functional analysis of yeast centromere DNAs.** *Cell* 1982, **29**:235-244.
 13. Hyman AA, Middleton K, Centola M, Mitchison TJ, Carbon J: **Microtubule-motor activity of a yeast centromere-binding protein complex.** *Nature* 1992, **359**:533-536.
 14. Meluh PB, Yang P, Glowczewski L, Koshland D, Smith MM: **Cse4p is a component of the core centromere of *Saccharomyces cerevisiae*.** *Cell* 1998, **94**:607-613.
 15. Bernard P, Maure JF, Partridge JF, Genier S, Javerzat JP, Allshire RC: **Requirement of heterochromatin for cohesion at centromeres.** *Science* 2001, **294**:2539-2542.
 16. Blower MD, Sullivan BA, Karpen GH: **Conserved organization of centromeric chromatin in flies and humans.** *Dev Cell* 2002, **2**:319-330.
 17. Henikoff S, Ahmad K, Malik HS: **The centromere paradox: stable inheritance with rapidly evolving DNA.** *Science* 2001, **293**:1098-1102.
 18. Nekrasov VS, Smith MA, Peak-Chew S, Kilmartin JV: **Interactions between centromere complexes in *Saccharomyces cerevisiae*.** *Mol Biol Cell* 2003, **14**:4931-4946.
 19. De Wulf P, McAinsh AD, Sorger PK: **Hierarchical assembly of the budding yeast kinetochore from multiple subcomplexes.** *Genes Dev* 2003, **17**:2902-2921.
 20. Baldauf SL, Roger AJ, Wenk-Siefert I, Doolittle WF: **A kingdom-level phylogeny of eukaryotes based on combined protein data.** *Science* 2000, **290**:972-977.
 21. Sullivan KF: **A solid foundation: functional specialization of centromeric chromatin.** *Curr Opin Genet Dev* 2001, **11**:182-188.
 22. Hosouchi T, Kumekawa N, Tsuruoka H, Kotani H: **Physical map-based sizes of the centromeric regions of *Arabidopsis thaliana* chromosomes 1, 2, and 3.** *DNA Res* 2002, **9**:117-121.
 23. Sun X, Le HD, Wahlstrom JM, Karpen GH: **Sequence analysis of a functional *Drosophila* centromere.** *Genome Res* 2003, **13**:182-194.
- The best-characterized functional centromere in a key model organism is shown to contain a mixture of satellite repeats, transposon fragments and intact retrotransposons with no evidence of centromere-specific sequences. Sequencing of satellite-containing centromeric DNA remains difficult but, as this study shows, not impossible.
24. Rudd MK, Schueler MG, Willard HF: **Sequence organization and functional annotation of human centromeres.** *Cold Spring Harb Symp Quant Biol* 2003, **68**:141-149.
- Homogenous α -satellite repeats that show a higher order structure appear to be the major, if not the exclusive, sequence components of human centromeres, but they are still strongly under-represented in sequence databanks.
25. Schueler MG, Higgins AW, Rudd MK, Gustashaw K, Willard HF: **Genomic and genetic definition of a functional human centromere.** *Science* 2001, **294**:109-115.
 26. Kumekawa N, Hosouchi T, Tsuruoka H, Kotani H: **The size and sequence organization of the centromeric region of *Arabidopsis thaliana* chromosome 4.** *DNA Res* 2001, **8**:285-290.
 27. Moore LL, Roth MB: **HCP-4, a CENP-C-like protein in *Caenorhabditis elegans*, is required for resolution of sister centromeres.** *J Cell Biol* 2001, **153**:1199-1208.
 28. Howman EV, Fowler KJ, Newson AJ, Redward S, MacDonald AC, Kalitsis P, Choo KH: **Early disruption of centromeric chromatin organization in centromere protein A (CenpA) null mice.** *Proc Natl Acad Sci USA* 2000, **97**:1148-1153.
 29. Oegema K, Desai A, Rybina S, Kirkham M, Hyman AA: **Functional analysis of kinetochore assembly in *Caenorhabditis elegans*.** *J Cell Biol* 2001, **153**:1209-1226.
 30. Van Hooser AA, Ouspenski II, Gregson HC, Starr DA, Yen TJ, Goldberg ML, Yokomori K, Earnshaw WC, Sullivan KF, Brinkley BR: **Specification of kinetochore-forming chromatin by the histone H3 variant CENP-A.** *J Cell Sci* 2001, **114**:3529-3542.
 31. Westermann S, Cheeseman IM, Anderson S, Yates JR, Drubin DG, Barnes G: **Architecture of the budding yeast kinetochore reveals a conserved molecular core.** *J Cell Biol* 2003, **163**:215-222.
 32. Palmer DK, O'Day K, Wener MH, Andrews BS, Margolis RL: **A 17-kD centromere protein (CENP-A) copurifies with nucleosome core particles and with histones.** *J Cell Biol* 1987, **104**:805-815.
 33. Yoda K, Ando S, Morishita S, Houmura K, Hashimoto K, Takeyasu K, Okazaki T: **Human centromere protein A (CENP-A) can replace histone 3 in nucleosome reconstitution *in vitro*.** *Proc Natl Acad Sci USA* 2000, **97**:7266-7271.
 34. Black BE, Foltz DR, Chakravarthy S, Luger K, Woods VL Jr, Cleveland DW: **Structural determinants for generating centromeric chromatin.** *Nature* 2004, **430**:578-582.
- Physical measurements suggest that the shape of the CENP-A/H4 tetramer distinguishes it from the canonical tetramer. It will be interesting to see whether this difference is involved in discrimination between centromeric and canonical nucleosomes.
35. Vermaak D, Hayden HS, Henikoff S: **Centromere targeting element within the histone fold domain of Cid.** *Mol Cell Biol* 2002, **22**:7553-7561.
 36. Malik HS, Henikoff S: **Adaptive evolution of Cid, a centromere-specific histone in *Drosophila*.** *Genetics* 2001, **157**:1293-1298.
 37. Cooper JL, Henikoff S: **Adaptive evolution of the histone fold domain in centromeric histones.** *Mol Biol Evol* 2004, **21**:1712-1718.
 38. Talbert PB, Bryson TD, Henikoff S: **Adaptive evolution of centromere proteins in plants and animals.** *J Biol* 2004, **3**:18.
- Adaptive evolution reflects genetic conflict and now both major DNA-binding proteins at centromeres are found to be adaptively evolving. CENP-C is seen as the major suppressor of centromere meiotic drive, even in lineages such as mammals and grasses in which CenH3s are under strict purifying selection. This drive process could account for the complexity and rapid evolution of centromeres in plants and animals.
39. Wieland G, Orthaus S, Ohndorf S, Diekmann S, Hemmerich P: **Functional complementation of human centromere protein A (CENP-A) by Cse4p from *Saccharomyces cerevisiae*.** *Mol Cell Biol* 2004, **24**:6620-6630.
- Human and yeast centromeres lie at extreme ends of the complexity spectrum, yet the unique component of centromeric chromatin from yeast is not only targeted to human centromeres but is also fully functional in a human cell line. This finding would appear to limit models for centromere identity in humans to those that do not require sequence constraints in the process of centromeric nucleosome assembly.

40. Zeng K, de las Heras JI, Ross A, Yang J, Cooke H, Shen MH: **Localisation of centromeric proteins to a fraction of mouse minor satellite DNA on a mini-chromosome in human, mouse and chicken cells.** *Chromosoma* 2004, **113**:84-91.

41. Tyler-Smith C, Gimelli G, Giglio S, Florida G, Pandya A, Terzoli G, Warburton PE, Earnshaw WC, Zuffardi O: **Transmission of a fully functional human neocentromere through three generations.** *Am J Hum Genet* 1999, **64**:1440-1444.

42. Amor DJ, Bentley K, Ryan J, Perry J, Wong L, Slater H, Choo KH: **Human centromere repositioning "in progress".** *Proc Natl Acad Sci USA* 2004, **101**:6542-6547.

A centromere shift from the normal position to a distant euchromatic site on an unrearranged human chromosome 4 is not associated with any phenotypic abnormalities. The spontaneous emergence of a new centromere on an otherwise normal chromosome might represent the first stage in centromere evolution.

43. Ventura M, Weigl S, Carbone L, Cardone MF, Misceo D, Teti M, D'Addabbo P, Wandall A, Björck E, de Jong PJ *et al.*: **Recurrent sites for new centromere seeding.** *Genome Res* 2004, **14**:1696-1703.

Some karyotypic changes in the primate lineage are shown to be attributable to ancient neocentromere formation events. It is suggested that neocentromeres form in particular regions, although examples presented in this paper on human neocentromeres are consistent with those of reference [45] in which no preferential region for neocentromere formation was detected.

44. Saffery R, Sumer H, Hassan S, Wong LH, Craig JM, Todokoro K, Anderson M, Stafford A, Andy Choo KH: **Transcription within a functional human centromere.** *Mol Cell* 2003, **12**:509-516.

Human neocentromeres form in apparently ordinary regions that contain active genes. No effect on gene expression was detected when comparing a neocentromere with the parental normal chromosome.

45. Nagaki K, Cheng Z, Ouyang S, Talbert PB, Kim M, Jones KM, Henikoff S, Buell CR, Jiang J: **Sequencing of a rice centromere uncovers active genes.** *Nat Genet* 2004, **36**:138-145.

The first sequence of a centromere in a complex eukaryote was accomplished on a chromosome with only a small region of satellite DNA. The kinetochore region was mapped to a 750 kb span that contains several active genes, is rich in methylation of histone H3 Lys9, is poor in H3 Lys4 and appears to show plasticity in chromatin composition.

46. Alonso A, Mahmood R, Li S, Cheung F, Yoda K, Warburton PE: **Genomic microarray analysis reveals distinct locations for the CENP-A binding domains in three human chromosome 13q32 neocentromeres.** *Hum Mol Genet* 2003, **12**:2711-2721.

By mapping three neocentromeres that arose independently in the same cytological position, the authors asked whether there are preferential positions for centromere formation. Surprisingly, all three neocentromeres were found to occur at different positions and, therefore, much of the genome is competent to form centromeres.

47. Wu J, Yamagata H, Hayashi-Tsugane M, Hijishita S, Fujisawa M, Shibata M, Ito Y, Nakamura M, Sakaguchi M, Yoshihara R *et al.*: **Composition and structure of the centromeric region of rice chromosome 8.** *Plant Cell* 2004, **16**:967-976.

48. Burke DT, Carle GF, Olson MV: **Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors.** *Science* 1987, **236**:806-812.

49. Ahmad K, Henikoff S: **Centromeres are specialized replication domains in heterochromatin.** *J Cell Biol* 2001, **153**:101-110.

50. Shelby RD, Monier K, Sullivan KF: **Chromatin assembly at kinetochores is uncoupled from DNA replication.** *J Cell Biol* 2000, **151**:1113-1118.

51. Ahmad K, Henikoff S: **Histone H3 variants specify modes of chromatin assembly.** *Proc Natl Acad Sci USA* 2002, **99**(Suppl 4):16477-16484.

52. Shelby RD, Vafa O, Sullivan KF: **Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites.** *J Cell Biol* 1997, **136**:501-513.